Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 4267–4270, June 1988
Biochemistry

Stress proteins are immune targets in leprosy and tuberculosis

(heat shock/*dnaK/groEL*/vaccines)

DOUGLAS YOUNG*, RAJU LATHIGRA*, ROGER HENDRIX†, DOUG SWEETSER‡, AND RICHARD A. YOUNG‡

*Medical Research Council Tuberculosis and Related Infections Unit, Hammersmith Hospital, London W12 0HS, United Kingdom; †Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260; and ‡Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Barry R. Bloom, March 7, 1988

ABSTRACT **To understand the immune response to infection by tuberculosis and leprosy bacilli and to develop improved vaccines, the nature of antigens that are involved in humoral and cell-mediated immunity was investigated. We have determined that five immunodominant protein antigens under study are homologues of stress proteins. This finding and observations with other pathogens suggest that infectious agents may respond to the host environment by producing stress proteins and that these proteins can be important immune targets. We postulate that abundant and highly conserved stress proteins may have "immunoprophylactic" potential for a broad spectrum of human pathogens.**

Mycobacterium tuberculosis and *Mycobacterium leprae* are the etiologic agents of tuberculosis and leprosy, respectively. These diseases afflict 20–30 million people and continue to present a significant global health problem (1, 2). To develop more effective tools for the diagnosis and prevention of these diseases, it is important to understand the immune response to infection by mycobacterial pathogens.

The antibody and T-cell responses to infection or inoculation with killed mycobacteria have been studied in humans and in animals. Human patients with tuberculosis or leprosy produce serum antibodies directed against at least 12 mycobacterial proteins. Some of these proteins are also recognized by well-characterized murine monoclonal antibodies. Mice immunized with mycobacterial lysates produce antibodies that are directed predominantly to six *M. tuberculosis* and six *M. leprae* protein antigens (3, 4). Genes encoding these 12 mycobacterial antigens have been cloned (5–7), and recombinant proteins produced from these clones have been used to investigate the human T-lymphocyte response to mycobacterial infection.

Protection against mycobacterial disease involves cell-mediated immunity (1, 8). T lymphocytes cloned from patients or from volunteers immunized with killed mycobacteria have been tested for their ability to recognize the recombinant mycobacterial proteins. Lymphocyte-proliferation assays demonstrate that most of the antigens identified with monoclonal antibodies are involved in the T-cell response to mycobacterial infection or vaccination in mice (9, 10) and in humans (11–20). Limiting dilution analysis indicates that 20% of the mycobacterial-reactive $CD4^+$ T lymphocytes in mice immunized with *M. tuberculosis* recognize a single protein, the 65-kDa antigen (9).

In view of the involvement of these proteins in humoral and cell-mediated immune responses, we wished to obtain clues to the functions of these proteins in the mycobacterial cell and to sequence the DNA encoding several of the *M. tuberculosis* and *M. leprae* antigens.§ We report here that many of these mycobacterial protein antigens exhibit striking sequence similarity to known stress-induced proteins and discuss the implications of this observation.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Recombinant DNA Clones. The isolation and characterization of *M. tuberculosis* and *M. leprae* λgt11 genomic DNA clones with murine monoclonal antibodies have been described (5, 6). DNA was isolated from these clones and was manipulated by standard procedures (21).

DNA Sequence Analysis. DNA was subcloned into vector M13mp18 or M13mp19 (New England Biolabs) as suggested by the supplier. Dideoxynucleotide chain-termination reactions and gel electrophoresis of the sequenced products were as described (22). DNA sequences were determined for both strands of DNA. Computer analysis of sequences with UWGCG programs was as described by Devereux *et al.* (23).

Immunoblot Analysis. *Escherichia coli* strain TG1 was transformed with the following plasmids by standard procedures (24) with selection for ampicillin resistance: pND5, a derivative of pBR325 containing the *E. coli groE* genes (25); pUC8 (26); pUC8 with the insert DNA from λgt11 clone Y3178 (*M. leprae* 65-kDa antigen, ref. 6) ligated in the *Eco*RI site.

Overnight cultures of *E. coli* strains in Luria–Bertani (LB) medium were centrifuged and resuspended in isotonic phosphate-buffered saline at a cell density corresponding to an absorbance of 20 at 600 nm. An equal volume of sample buffer containing 2% (wt/vol) $NaDodSO_4$ was added, and, after heating on a boiling water bath for 2 min, 5-μl samples were electrophoresed on 12% (wt/vol) polyacrylamide gels in the presence of $NaDodSO_4$. Blots were prepared by electrophoretic transfer of the proteins to a nitrocellulose membrane, and binding of monoclonal antibodies was assayed with a peroxidase-conjugated secondary antibody as described (27).

RESULTS AND DISCUSSION

Six *M. tuberculosis* and six *M. leprae* proteins have been implicated in the immune response to the mycobacterial pathogens (Table 1). To obtain clues to the normal cellular function of several of these mycobacterial antigens, DNA clones encoding these proteins, isolated by using monoclonal antibodies to probe λgt11 libraries (5, 6), were subjected to sequence analysis. The sequences elucidated have been submitted to the GenBank sequence database.§

The Mycobacterial 71-kDa Antigen. The 71-kDa antigen of *M. tuberculosis* is recognized by human T cells during infection (Table 1). The insert DNA of λgt11 clone Y3271 (5)

§The sequences for these proteins have been deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03838–J03840).

Table 1. Mycobacterial protein antigens

Protein, kDa	Recognized by human T cells	Subjected to sequence analysis	Homology with known proteins	Ref(s).
M. tuberculosis				
71	+	+	DnaK	16
65*	+	+	GroEL	17, 19
38	+	–	–	15
19	+	+	None	19
14	+	–	–	—†
12	ND	–	–	
M. leprae				
70	ND	–	DnaK	
65	+	+	GroEL	18
36	+	–	–	12
28	+	–	–	—†
18	+	+	Plant Hsp	11
12	ND	–	–	12

Mycobacterial protein antigens, their recognition by human T cells, and homology of the deduced mycobacterial protein sequences to known proteins are summarized. ND, not determined. +, Yes; –, no.

*Includes data derived from study of the 65-kDa antigen of *M. bovis* BCG (bacillus Calmette–Guérin), which is identical to the *M. tuberculosis* 65-kDa antigen.

†A. S. Mustafa, J. R. Lamb, D.Y., and R.A.Y., unpublished data.

was sequenced to obtain amino acid sequence information for the 71-kDa antigen of *M. tuberculosis*. This clone produces a β-galactosidase fusion protein containing the carboxyl-terminal one-third of the 71-kDa antigen. A comparison of the sequence obtained for the carboxyl-terminal one-third of the 71-kDa antigen with those in the GenBank database[¶] revealed that the *M. tuberculosis* 71-kDa antigen exhibits 40% amino acid sequence identity with the comparable segment of the *dnaK* gene product from *E. coli* (28) (Fig. 1). Fig. 2*A* shows the extent of sequence similarity between portions of the mycobacterial and the *E. coli* 70-kDa polypeptides. Sequences transcriptionally downstream from the mycobacterial 71-kDa gene predict a 356-amino acid protein homologous to the *E. coli dnaJ* gene product (unpublished data), indicating that the *E. coli dnaK–dnaJ* operon structure is conserved in *M. tuberculosis* and consistent with the conclusion that the mycobacterial 71-kDa antigen is a homologue of the *E. coli dnaK* gene product. The product of the *dnaK* gene is a member of the 70-kDa heat shock protein family that is highly conserved among prokaryotes and eukaryotes (28, 29).

The *M. leprae* 70-kDa antigen cross-reacts with monoclonal antibodies directed to the *M. tuberculosis* 71-kDa protein (27). The similarity in size and antigenicity indicate that these two mycobacterial proteins are homologues. Thus, we conclude that the 70-kDa antigens of *M. tuberculosis* and *M. leprae* are both members of the 70-kDa heat shock protein family of stress proteins.

The Mycobacterial 65-kDa Antigen. The 65-kDa antigens of *M. tuberculosis* and *M. leprae* are involved in the human T-cell response to mycobacterial infection (Table 1). Genes encoding these proteins have been isolated (5, 6) and sequenced (30, 31), revealing that the amino acid sequences of the 65-kDa antigens of *M. tuberculosis* and *M. leprae* are 95% identical. These protein sequences exhibit no significant sequence similarity to proteins in the GenBank database.[¶]

To identify these proteins, we exploited the observation that some monoclonal antibodies directed against the mycobacterial 65-kDa antigens cross-react with an *E. coli* protein of 60 kDa (Fig. 3). We observed that *E. coli* cells transformed with the plasmid pND5 (22), which contains the *E. coli groE* genes, accumulate large amounts of the 60-kDa protein (Fig. 3). A comparison of the mycobacterial 65-kDa protein sequences with those determined for *E. coli groEL* (C. Woolford, K. Tilly, C. Georgopoulos, and R.H., unpublished data) revealed 60% amino acid sequence identity. The extent of this sequence similarity is shown in Fig. 2*B*.

The 60-kDa GroEL protein is a major stress protein in *E. coli* (29). There is some evidence that the mycobacterial 65-kDa proteins accumulate in response to stress; *Mycobacterium bovis* BCG (bacillus Calmette–Guérin) cultures grown in zinc-deficient medium are substantially enriched in this protein (14). We infer that the 65-kDa proteins of *M. tuberculosis* and *M. leprae* are homologues of the *E. coli* GroEL protein.

Other Mycobacterial Antigens. T lymphocytes that respond to the *M. tuberculosis* 19-kDa antigen and the *M. leprae* 18-kDa antigen have been observed in humans with tuberculosis and leprosy, respectively (Table 1). DNA encoding these antigens was sequenced from the λgt11 clonès Y3148 (5) and Y3179 (6), respectively. The *M. tuberculosis* 19-kDa protein sequence predicted from the DNA exhibited no significant sequence similarity to proteins in the GenBank database.[¶] However, the *M. leprae* 18-kDa protein sequence was similar to the soybean 17-kDa heat shock protein, a protein representative of a major class of plant heat shock proteins (ref. 32; A. Nerland and D.S., unpublished data).

Implications. We find that three of the *M. leprae* protein antigens and two of the *M. tuberculosis* protein antigens under study exhibit striking sequence similarity to known stress proteins. For reasons expressed above, we conclude that two of the *M. leprae* and two of the *M. tuberculosis* antigens are homologues of the *E. coli* DnaK and GroEL proteins. What is the relationship between stress proteins and the host immune response to mycobacterial infection? When cells are subjected to a variety of stresses, they respond by selectively increasing the synthesis of a limited set of stress

[¶]EMBL/GenBank Genetic Sequence Database (1986) GenBank (IntelliGenetics, Mountain View, CA), Tape Release 46.0.

```
E F Q P S V Q I Q V Y Q G E R E I A A H N K L L G   25
D N * G A * T * H * L * * * * K R * * D * * S * *

S F E L T G I P P A P R G I P Q I E V T F D I D A   50
Q * N * D * * N * * * * * N * * * * * * * * * * *

N G I V H V T A K D K G T G K E N T I R I Q E G S   75
D * * L * * S * * * * N S * * * Q K * T * K A S *

G L S K E D I D R M I K D A E A H A E E D R K R R  100
* * N E D E * Q K * V R * * * * N * * A * * * F E

E E A D V R N Q A E T L V Y N T E K F V K E Q R E  125
* L V Q T * * * G D H * L H S * R * Q * E * A G D

G G S K V P E D T W R I G Y F G H Q V G D G E A G  150
K L P A D D K T A I E S A L T A L E T A L K G E D

P G V A G S G A S D L R S S S G C V T G H W R C P  175
K A A I E A K M Q E * A Q V * Q K L M E I A Q Q Q

P R A A A G R C P P R L G M. tuberculosis 71 kDa
H A Q Q Q T A G A D A S A E. coli DnaK (431-618)
```

FIG. 1. Comparison of *M. tuberculosis* 71-kDa antigen and *E. coli* DnaK protein sequences. The sequence of the carboxyl-terminal 363 amino acids of the 71-kDa antigen (top line) is aligned with the corresponding region of the DnaK protein from *E. coli* (28). Identical residues are indicated by a star.

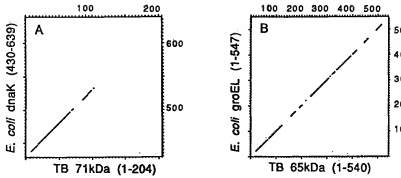

FIG. 2. Homologies between mycobacterial antigens and known stress proteins. Amino acid sequences of mycobacteria and *E. coli* proteins were compared by using a dot matrix program (23) with the window set at 12 and the stringency at 6. The matrices indicate sequence similarity between portions of the *M. tuberculosis* 71-kDa antigen (residues 1–204; TB 71 kDa) and the *E. coli* DnaK protein (residues 430–639) (*A*) and the *M. tuberculosis* 65-kDa antigen (residues 1–540; TB 65 kDa) and the *E. coli* GroEL protein (residues 1–547) (*B*).

proteins. Some stress proteins, including the products of *dnaK* and *groEL*, are major constituents of the cell under normal growth conditions and are induced to even higher levels during stress (29, 33). In retrospect, the fact that stress-related proteins are targets of the immune response could have been anticipated. One might expect that immunodominant antigens would be found among such abundant proteins.

Although the function of stress proteins is not entirely clear, it appears that some participate in assembly and structural stabilization of certain cellular and viral proteins (33–36), and their presence at high concentrations may have an additional stabilizing effect during exposure to adverse conditions. Phagocytic host cells produce a hostile environment for foreign organisms, and the ability to produce stress proteins has been implicated in the survival of bacterial pathogens within macrophages (37, 38). We postulate that these abundant proteins are then among the major antigens available for presentation to T lymphocytes and that this may contribute to their antigenicity.

Stress proteins may be common immune targets in a broad spectrum of infectious diseases. Sequence analysis has revealed 70-kDa heat shock protein homologues among major antigens of the protozoan parasites *Plasmodium falciparum* (39) and *Schistosoma mansoni* (40) and the filarial parasite *Brugia malayi* (41). Similarly, homologues of GroEL have been found among antigens involved in the immune response to *Salmonella typhimurium* and *Coxiella* (42). The presence of stress proteins among major immune targets in a variety of human pathogens suggests that the stress response may be a general component of infection and that stress proteins should be considered among candidates for subunit vaccines. Moreover, because the stress response is common to prokaryotes and eukaryotes and some stress proteins are highly conserved in sequence (28, 29, 43), it is possible that an antigen from one pathogen could immunize against another pathogen. Exposure to foreign stress proteins early in life might, in fact, induce a degree of immunity to a variety of infectious agents. If so, the presence of this immunity could provide an explanation for the observation that, for many pathogens, only a fraction of infected individuals actually acquire clinical disease.

Immunization with conserved protein antigens, whether during natural infection or during vaccination, might also have negative consequences. Presentation of a self-like determinant may fail to stimulate lymphocytes due to host tolerance. Chronic presentation of such determinants might occasionally result in breaking of tolerance and production of an autoimmune response.

We have observed an intriguing relationship between stress proteins and the immune response to mycobacterial infection. A more detailed examination of stress protein determinants and immune response mechanisms is essential to understanding the relationships among stress proteins, infection, and immunity.

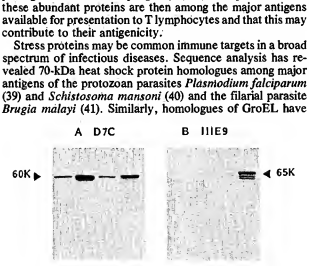

FIG. 3. Antigenic cross-reactivity between the 65-kDa antigen of mycobacteria and *E. coli* GroEL. Protein extracts from *E. coli* strain TG1 (lanes 1), TG1 plus a plasmid containing the *groEL* and *groES* genes (pND5) (lanes 2), TG1 plus pUC8 (lanes 3), or TG1 plus a pUC8 derivative containing the insert DNA from the λgt11 clone Y3178 (6) expressing the 65-kDa antigen of *M. leprae* (lanes 4) were subjected to $NaDodSO_4$/polyacrylamide gel electrophoresis and were electrophoretically transferred to a nitrocellulose membrane. Duplicate blots were probed with the monoclonal antibody D7C (*A*), which recognizes a 65-kDa epitope shared by a variety of bacteria, and the monoclonal antibody MLIIIE9 (*B*), which recognizes a 65-kDa epitope specific to *M. leprae*. K, kDa.

We thank Barry Bloom, Juraj Ivanyi, and Jonathan Lamb for helpful discussion; Millicent Masters for plasmid pND5; and Tom Buchanan for monoclonal antibodies. This work was supported by grants from the National Institutes of Health (AI23545), the World Health Organization Program for Vaccine Development, and the World Health Organization/World Bank/United Nations Development Program Special Program for Research and Training in Tropical Diseases. R.A.Y. is a Burroughs Wellcome Scholar in Molecular Parasitology.

1. Joint International Union Against Tuberculosis and World Health Organization Study Group (1982) *Tubercle* **63**, 157–169.
2. Bloom, B. & Godal, T. (1983) *Rev. Infect. Dis.* **5**, 765–780.
3. Engers, H. D., Abe, M., Bloom, B. R., Mehra, V., Britton, W., Buchanan, T. M., Khanolkar, S. K., Young, D. B., Closs, O., Gillis, T., Harboe, M., Ivanyi, J., Kolk, A. H. J. & Shepard, C. C. (1985) *Infect. Immun.* **48**, 603–605.
4. Engers, H. D., Houba, V., Bennedsen, J., Buchanan, T. M., Chaparas, S. D., Kadival, G., Closs, O., David, J. R., van Embden, J. D. A., Godal, T., Mustafa, A. S., Ivanyi, J., Young, D. B., Kaufmann, S. K. E., Khomenko, A. G., Kolk, A. H. J., Kubin, M., Louis, J. A., Minden, P., Shinnick, T., Trnka, L. & Young, R. A. (1986) *Infect. Immun.* **51**, 718–720.

5. Husson, R. N. & Young, R. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1679–1683.
6. Young, R. A., Mehra, V., Sweetser, D., Buchanan, T. M., Clark-Curtiss, J., Davis, R. W. & Bloom, B. R. (1985) *Nature (London)* **316**, 450–452.
7. Britton, W. J., Garsia, R. J., Hellqvist, L., Watson, J. D. & Basten, A. (1986) *Lepr. Rev.* **57**, Suppl. 2, 67–75.
8. Hahn, H. & Kaufmann, S. H. E. (1981) *Rev. Infect. Dis.* **3**, 1221–1250.
9. Kaufmann, S. H. E., Vath, U., Thole, J. E. R., van Embden, J. D. A. & Emmrich, F. (1987) *Eur. J. Immunol.* **17**, 351–357.
10. Boom, W. H., Husson, R. N., Young, R. A., David, J. R. & Piessens, W. F. (1987) *Infect. Immun.* **55**, 2223–2229.
11. Mustafa, A. S., Gill, H. K., Nerland, A., Britton, W. J., Mehra, V., Bloom, B. R., Young, R. A. & Godal, T. (1986) *Nature (London)* **319**, 63–66.
12. Ottenhoff, T. H. M., Klatser, P. R., Ivanyi, J., Elferink, D. G., de Wit, M. Y. L. & de Vries, R. R. P. (1986) *Nature (London)* **319**, 66–68.
13. Lamb, J. R. & Young, D. B. (1987) *Immunology* **60**, 1–5.
14. De Bruyn, J., Bosmans, R., Turneer, M., Weckx, M., Nyabenda, J., van Vooren, J.-P., Falmagne, P., Wiker, H. G. & Harboe, M. (1987) *Infect. Immun.* **55**, 245–252.
15. Young, D. B., Kent, L., Rees, A., Lamb, J. & Ivanyi, J. (1986) *Infect. Immun.* **54**, 177–183.
16. Britton, W. J., Hellqvist, L., Basten, A. & Inglis, A. S. (1986) *J. Exp. Med.* **164**, 695–708.
17. Emmrich, F., Thole, J., van Embden, J. & Kaufmann, S. H. E. (1986) *J. Exp. Med.* **163**, 1024–1029.
18. Lamb, J. R., Ivanyi, J., Rees, A. D. M., Rothbard, J. B., Holland, K., Young, R. A. & Young, D. B. (1987) *EMBO J.* **6**, 1245–1249.
19. Oftung, F., Mustafa, A. S., Husson, R., Young, R. A. & Godal, T. (1987) *J. Immunol.* **138**, 927–931.
20. Lamb, J. R., Ivanyi, J., Rees, A., Young, R. A. & Young, D. B. (1986) *Lepr. Rev.* **57**, Suppl. 2, 131–137.
21. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics: A Manual for Genetic Engineering* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
22. Sanger, F., Nicklén, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
25. Jenkins, A. J., March, J. B., Oliver, I. R. & Masters, M. (1986) *Mol. Gen. Genet.* **202**, 446–454.
26. Vieria, J. & Messing, J. (1982) *Gene* **19**, 259–268.
27. Young, D. B., Kent, L. & Young, R. A. (1987) *Infect. Immun.* **55**, 1421–1425.
28. Bardwell, J. C. A. & Craig, E. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 848–852.
29. Lindquist, S. (1986) *Annu. Rev. Biochem.* **55**, 1151–1191.
30. Shinnick, T. M. (1987) *J. Bacteriol.* **169**, 1080–1088.
31. Mehra, V., Sweetser, D. & Young, R. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7013–7017.
32. Schoffl, F., Raschke, E. & Nagac, R. T. (1984) *EMBO J.* **3**, 2491–2497.
33. Neidhardt, F. C. & VanBogelen, R. A. (1987) in *Escherichia coli and Salmonella typhimurium*, Cellular and Molecular Biology, eds. Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E. (Am. Soc. Microbiol., Washington, DC), pp. 1334–1345.
34. Pelham, H. R. B. (1986) *Cell* **46**, 959–961.
35. Takano, T. & Kakefuda, T. (1972) *Nature (London) New Biol.* **239**, 34–37.
36. Georgopoulos, C., Hendrix, R., Kaiser, A. D. & Wood, W. B. (1972) *Nature (London) New Biol.* **239**, 38–41.
37. Christman, M. F., Morgan, R. W., Jacobson, F. S. & Ames, B. N. (1985) *Cell* **41**, 753–762.
38. Morgan, R. W., Christman, M. F., Jacobson, F. S. & Ames, B. N. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8059–8063.
39. Bianco, A. E., Favaloro, J. M., Burkot, T. R., Culvenor, J. G., Crewther, P. E., Brown, G. V., Anders, R. F., Coppel, R. L. & Kemp, J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8713–8717.
40. Hedstrom, R., Culpepper, J., Harrison, R. A., Agabian, N. & Newport, G. (1987) *J. Exp. Med.* **165**, 1430–1435.
41. Selkirk, M. E., Rutherford, P. J., Denham, D. A., Partono, F. & Maizels, R. M. (1988) *Biochem. Soc. Symp.*, in press.
42. Wodkin, M. H. & Williams, J. C. (1988) *J. Bacteriol.*, in press.
43. Bardwell, J. C. A. & Craig, E. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5177–5181.